SYSTEM AND METHOD FOR IDENTIFYING CLOAKED WEB SERVERS

ABSTRACT OF THE DISCLOSURE

A search engine receives from a client a representation of a first object that was returned by a web server to the client in response to a request from the client. The search engine receives from the web server a second object in response to an identical request from the search engine, and compares the representation of the first object to a representation of the second object. The web server is determined to be cloaked if the representation of the first object does not match the representation of the second object. Typically, the client receives a URL embedded in a response to a search request submitted to the search engine. A toolbar operating in conjunction with the web browser on the client processes the URL. The processing includes: directing the web browser to obtain an object corresponding to the URL from a web server addressed by the URL; converting the object to a feature vector; and delivering the feature vector and the URL back to the search engine.